DOCUMENT RESUME

ED 337 746 CS 010 714

AUTHOR Griffin, Patrick
TITLE Literacy Assessment: Merging Teaching, Learning, and Assessment.
PUB DATE May 91
NOTE 13p.; Paper presented at the Annual Meeting of the International Reading Association (36th, Las Vegas, NV, May 6-10, 1991).
PUB TYPE Speeches/Conference Papers (150) -- Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Criterion Referenced Tests; *Educational Assessment; Elementary Secondary Education; *Literacy; *Standardized Tests; Student Evaluation; *Test Use; Test Validity
IDENTIFIERS *Teacher Judgments

ABSTRACT

Data from research studies seem to support the validity of teacher judgments of student achievement; however, there are unresolved issues, including the focus of the judgment, the external validity or credibility of the assessment, and the process of the assessment itself. To achieve maximum benefits from this form of assessment/teaching interface, teachers may need to be sensitized to the extent and importance of the assessment role in the teaching process. Intensive experience with basic principles of measurement and assessment instruments and other devices is needed. Effective teaching is based on an ability to assess accurately: assessment does and should inform the teaching process. It becomes more powerful as the assessment becomes integrated into teaching. The more a teacher uses intuitive judgement and direct observation of a student's literacy development, the more powerful the assessment becomes in the classroom. The key to the development of a cohesive and coherent set of assessment tasks all directed toward the same developmental aim is the existence of a criterion scale of development that can be used to link the teaching and learning process and the interpretation of assessment information. Linking assessment and reporting to teaching and learning is facilitated by the use of descriptive criterion scales which standardize the reporting terminology. The link between teaching, learning, and assessment, while having an undeniable internal validity, may at times need an external reference to establish external validity. (Two figures are included; 27 references are attached.) (RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED337746

Literacy Assessment: Merging Teaching, Learning, and Assessment

Patrick Griffin

ASSESSMENT RESEARCH CENTRE
Phillip Institute of Technology

CS010714

Paper Presented at the 36th annual convention of the International Reading Association, Las Vegas, May 6-10, 1991.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Patrick Griffin

TO THE EDUCATIONAL RESOURCES INFORMATION

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

In this paper we will examine the possibilities of integrating assessment with teaching and learning. In doing so we will examine the interactive process of question and answer in the classroom and how it closely parallels a more formal process of assessment and reporting.

There is a growing body of evidence that indicates how teacher judgements influence decisions in the interactive phase of Teaching (McNair, 1978, Hoge and Colardarci, 1989). The teachers main consideration during reading instruction appears to be reading achievement. Teachers pace whole class instruction on the basis of whether an identifiable group of students understand what is being presented (Clark and Peterson, 1986).

Questions are used in the evaluation of pupil comprehension, learning, thinking, knowledge or task performance. (Colker, 1984). All this supports the idea that teacher decision-making, particularly in an interactive context, is influenced by judgements about student learning.

The issue is whether the teacher judgement is accurate. The implications are important when judgements inform decisions regarding students for feedback, reporting to parents and other stakeholders (Elliott, Gresham, Freeman and McCloskey, 1988). Teacher judgements provide the primary data for most classroom decisions. It is widely assumed that teachers are generally poor judges of student attributes and that this is due to lack of perception, bias and error. "Directly or indirectly, the accuracy of teacher's assessments of student ability is often an issue in educational research. It is commonly argued that commercial tests provide teachers with valuable information about the abilities and deficiencies of their students, from which it follows that teachers who rate their students without such information will be in error." (Egan and Archer, 1985, p.25)

Hoge and Colardarci however, through meta analysis of a range of studies, have concluded that teachers differ in how accurately they judge their students' achievement. There was a generally high level of agreement between judgemental measures and standardised achievement test scores. A median correlation of 0.62 was found. Griffin (1990) replicated this level of agreement when teachers were provided with a descriptive criterion scale on which to have the judgement of achievement.

Data from research studies seem to support the validity of teacher judgements of student achievement. Studies of both convergent and concurrent validity have been consistently higher than those reported for psychological tests, (Hoge and Colardarci, 1989), despite wide variations in methodology across studies. However there are still some unresolved issues. These include the focus of the judgement, the external validity or credibility of the assessment, and the process of the assessment itself.

The Focus:

Judgemental assessments do not always make it clear as to what aspect of student performance is being assessed. The work of Griffin (1990) and of Farr and Farr (1991) offer some guide in that both provide descriptive scales as a frame of reference for the judgements, and illustrate how the judgement is closely allied to the teaching and learning process. In providing the descriptions of increasing proficiency, the materials circumvent the issue of the validity of standardised test data. Moreover, problems associated with global judgements (high, low, etc.) are avoided or at least contracted by reference to the descriptive scales.

Coladarci's (1986) criticism is avoided because much is disclosed about the teachers' specific knowledge of what the student has and has not mastered in some domain: In the case of the Griffin (1990) and Farr and Farr (1991) projects, the domains are Reading and Writing respectively. As yet, however, too little research has been done on convergent and discriminant validity of judgements based on criterion scales.

In order to achieve maximum benefits from this form of assessment/teaching interface however, teachers may need to be sensitised to the extent and importance of the assessment role in the teaching process (Hoge, 1983 and Hoge and Cudmore, 1986). Intensive experience with basic principles of measurement and assessment instruments and other devices, including norm referenced tests, observational procedures, and judgemental scales, are needed. Many more judgement scales need to be developed in line with Glaser's (1981) criterion referenced interpretation. Using this experience and improved assessment technology, there is a need to enhance teachers' abilities at analysing and diagnosing learning in children. A great deal of importance is attached to the judgements of teachers in both the teaching process and in communications with stakeholders in the education process.

The Nature of the Assessment:

Griffin & Nix (1991) define Assessment and Reporting as... *"the purposeful process of observing, interpreting, reporting and communicating information about student learning to relevant stakeholders"* (p.10). We can take each part of this definition and examine it for its relevance to teaching and learning. Assessment and reporting, and teaching and learning, are both purposeful processes. Each has as its major purpose to improve, assist and encourage learning. They are processes that continue throughout the teaching and learning procedure in the classroom. Assessment and reporting is not an event. It is not, as is often believed, a single test at the end of term and a report that goes home to parents. Assessment and reporting is not a one-off thing. It is a continuous process that underpins most, if not all, human interaction. It is the process of gathering information through various forms of observation. Each time two people interact, exchange information, cooperate, negotiate, or work together in any way, assessment takes place.

Information is gathered and interpreted in order to make sense of the observations that are being made. Several interpretive frameworks are used. We interpret what we see in terms of what we expect to see. That is a norm referenced assessment. We interpret what we see in terms of the actual behaviour and our knowledge of that kind of behaviour. That is criterion referenced interpretation. Alternatively the person can interpret their own behaviour for us. This is Ipsative referenced interpretation. We need to make sense of information and observations through all three frames of reference and to record the meaningful information. Interpretations are recorded in a notebook, a folio, or sometimes it is recorded as an object, a product of a students work - a sample. Sometimes it is recorded in the observers memory. The last part of assessment and reporting definition involves the communication that takes place. If it is to parents, it may well be a via parent/teacher interview, a formal term report, or a detailed letter from the school. If it is to the student, it may be simple feedback about performance on a classroom task. If it is to the school district, it may be the aggregate scores of the class on a test, or description of the kinds of goals and targets set for the students.

However, assessment and reporting is not a strictly linear process from purpose through observation, interpretation, recording and communicating to a particular audience. Collecting assessment information never stops. A teacher's intuition, judgement and observation skills are constantly being used. It is interesting to note that standardised tests were originally introduced to compensate for poor teacher judgement. We now know a lot more about learning, and about reading in particular. We now understand that there are few standardised tests that can reflect the learning process as well as the development of reading skills. In addition teachers are more skilled in observing and interpreting children's reading behaviour, and a new form of standardised assessment is emerging that capitalises on teachers expert judgement. We can now recognise and use the fact that teachers are continuously and intuitively making judgements about student learning.

Inside the classroom the teacher is continuously reporting by feedback and what Black (1987) calls feed forward. Let's take an example of self assessment and include a report. In the figure below, the passage "When is a Word Not a Word" (Hopkins, 1984) is reproduced. Take a few minutes to read the passage. Monitor what you are doing when reading on, reading back, and note where you use syntactic, semantic, contextual, and graphophonic cues. Where do you risk being wrong, omitting something all together? What do you use to make meaning? Is your guess consistent with your knowledge of words? Do you learn anything?

Place Hopkins Figure Here.

When you have "read" the passage turn to your neighbour and report on your experience. While reading, you were assessing or gathering information about your own reading strategies (metacognition). You recorded these (mentally), interpreted them in terms of your own knowledge of reading (criterion referenced), and communicated your interpretation to a peer. This is assessment and reporting.

In teaching we observe what students do, make sense of it, keep records, make decisions, and act on it. The process is fundamentally the same. Effective teaching is based on an ability to assess accurately. Assessment does and should inform the teaching process. It becomes more powerful as the assessment becomes integrated into teaching. The more a teacher uses intuitive judgement and direct observation the more powerful the assessment becomes in the classroom. In fact the more we investigate the assessment and reporting process the more it sounds like teaching and learning. Goals are set, negotiated, explored, and a purpose is established. Throughout the delivery of instruction, teachers and students are continuously observing. Teacher observing student, student observing teacher, student observing student. In this interactive process, information is exchanged, interpreted, and recorded. The teacher gathers information about the students, the students gather information about the teacher, about each other, and about the subject matter. They use their existing knowledge base to make sense of their observations and integrate it, thus expanding that knowledge base. This becomes a recording mechanism. On the basis of formal and informal records and interpretations, the teacher makes decisions about what steps to take next. The teacher continuously provides feedback.

Take the simple question and answer process. There is always a purpose. The teacher asks a question, receives an answer, interprets that answer in terms of an expected response, and in terms of the knowledge, skills and attitudes that are expected to be developed. A record is made instantly, at least in the mind of the teacher, and some feedback is given to the student. More than likely some feedback may come from the student to the teacher. Hence, in simple classroom interaction the formal process of assessment and reporting is replicated.

Validating the Process:

What we want to do now is to examine how it is best done, take a look at some examples in the language curriculum, and explore possibilities of how these examples might be linked together.

In both teaching and learning, and assessment and reporting, we need to understand what is expected to occur. In some forms of assessment this is often hidden or taken for granted in a series of tests, exercises or work samples. But, if these tests, work samples, and exercises, were collected together to form a cohesive and coherent set of assessment tasks all directed toward the same developmental aim, we have a powerful eliciting and interpretative framework for a teacher to make sense of observations in the classroom. The key to this interpretative framework is the existence of a criterion scale of development that can be used to link the teaching and learning process and the interpretation of the assessment information. Linking assessment and reporting to teaching and learning is facilitated by the use of descriptive criterion scales which standardise the reporting terminology. They free the teacher to assess and teach in a manner most suitable to the particular class while reporting within a common frame of reference. There are several examples of such a model. The Literacy Profiles exemplified in Figure 2, chart the development of reading and writing and provide advice to teachers about the kinds of classroom exercises that can be used to elicit the behaviours included in the scales. In a remarkably similar scale, the United Kingdom National Curriculum, targets and levels for reading, writing, speaking, and listening describe increasingly proficient literacy behaviour.

Place Figure 2 Here.
Pages from Teacher's handbook.

Another set of scales has been developed by Griffin & Forwood, (1991). In the United States, the NAEP scales, although derived from test items, provide the same kind of interpretative framework. The College Board (1986) developed the testing system, the Degrees of Reading power, which provides a different kind of descriptive scale. This scale describes reading development in terms of the kinds of reading material that can be handled rather than in terms of the reading skills being developed. Farr and Farr (1991) have produced a series of writing criterion scales, tied to set reading passages.

These examples all have in common an underlying criterion scale. The Degrees of Reading Power scale has been included because, while this is a set of cloze passage reading test, its scores are interpreted in terms of the kinds of reading material that a student can be expected to demonstrate or master. The scales make it easy for a teacher to monitor a student's

progress almost by direct observation of the kinds of reading material, the kinds of reading behaviour, attitudes, and reading activity within the classroom.

The teacher is guided in each of these schemes as to the kinds of activities that are most likely to illicit reading or writing behaviour described in the scales. Once the teacher begins to develop reading activities, which will develop and illicit the indicative behaviours included in the scale, assessment and teaching and learning become indistinguishable.

This is a very different aspect to the washback effect of testing. It is not teaching to a test; it is teaching and learning deliberately designed to illicit indicative behaviours which indicate developing proficiency in language skills. It is only when the teacher recognises these behaviours that they can conclude that the student is developing and moving forward on the descriptive criterion scales. Classroom exercises to encourage learning to occur, and direct observation is used to determine whether the indicative behaviours are present. The observations may then be directly interpreted in terms of the location of those descriptors on the scales. This is criterion referenced interpretation of assessment information. It is assessment of the direct behaviour being demonstrated and the relationship of that behaviour to other indicative behaviours, which together, form a cohesive description of developing proficiency.

The criterion scales can legitimately drive the curriculum by enabling the teacher to set realistic goals for the student and negotiate regarding the kind of classroom activities to be undertaken in order to progress along the scale. The more we begin to work with descriptive criterion scales, the more difficult it becomes to separate teaching and learning from assessment. The process of observing, interpreting, recording, and communicating becomes teaching, and teachers can base their work on a cohesive developmental progression for the students. It allows the teacher to make explicit the intuitive process of assessment in interaction with students. The teacher/student relationship can be articulated through the activity in the classroom and the relationship of those activities to the developmental proficiency in the language skills. It gives the assessment, and teaching and learning, internal validity within the closed system of the classroom. The assessment being so solidly linked to a teaching learning process, and in fact, indistinguishable from it. This means that the assessment data has unarguable internal validity.

However, there is a note of caution. The internal validity of the classroom process must be matched by its credibility to external audiences. The assessment and the reporting process must have external validity. Outside observers also have expectations and may also have an understanding of the kinds of behaviours that link and act together to form a cohesive and coherent definition of developing proficiency. External observers may also have a vested interest in the way in which progress in learning is demonstrated. To maintain credibility, teachers' assessments must also have external validity.

It is important that the assessment information and its interpretation should be able to be generalised to audiences, other than, but including the student, teacher, and parent. The communication needs to be in a form that makes it understandable, accurate and acceptable. It is an unfortunate fact that the teachers judgement is still not always accepted as accurate. As a consequence, the link between teaching, and learning, and assessment, while having undeniable internal validity, may at times need an external reference in order to establish

external validity. The external reference is sometimes in a form of standard assessment tasks that are acceptable to the teacher and the student, and can be part of the teaching assessment and learning process. They should provide a powerful form of communication to outside audiences, and establish the external validity of the assessment. The use of descriptive scales can become a method of bridging the gap from the internal validity of the classroom intuitive and judgement based assessment, to the external need for objectively based information that is considered to be credible and valid.

REFERENCES

Black, P.J. (1987), Report of the Task Group on Assessment and Testing. London: Department of Education and Science and Welsh Office.

Clark, C.M. and Peterson, P.L. (1986), Teachers' Thought Processes. In M.C. Wittrock (Ed.), *Third handbook of research on teaching*. New York: Macmillan. pp. 255-296.

Coladarci, T. (1986), *Accuracy of teacher judgments of student responses to standardized test items*. Journal of Educational Psychology. No. 78. pp. 141-146.

Colker, L. (April 1984), Teachers' interactive thoughts about pupil cognition. Paper presented at the annual meeting of the American Educational Research Association. Chicago.

College Board (1986a), *DRP handbook*. New York: College Entrance Examination Board.

Dawkins, J. (1990), The Language of Australia: Discussion paper on an Australian Literacy and Language Policy for the 1990's. Canberra: Department of Employment Education and Training.

Egan, O. and Archer, P. (1985), *The accuracy of teachers' ratings of ability: A regression model*. American Educational Research Journal. Vol. 22. pp. 25-34.

Elliott, S.N., Gresham, F.M., Freeman, T. and McCloskey, G. (1988), *Teacher and observer ratings of children's social skills: Validation of the Social Skills Rating Scales*. Journal of Psychoeducational Assessment. No. 6. pp. 152-161.

Estes, T.H. and Ester, J.J. (1989), Degrees of Reading Power as Virginia's Passport to Literacy. *Reading in Virginia*. No. 24, pp. 6-13.

Farr, R. and Farr, B. (1991), Integrated Assessment System: Technical Report. Sydney. San Antonio: The Psychological Corporation.

Glaser, R. (1981), *The Future of Testing: A Research Agenda for Cognitive Psychology and Psychometrics*, American Psychologist. Vol. 36, pp. 923-936.

Griffin, P. (1990), *Profiling literacy development: monitoring the accumulation of reading skills*. Australian Journal of Education. Vol. 34. No. 3. pp. 290-311.

Griffin, P. and Nix, P. (1991), Educational Assessment and Reporting: A New Approach. Sydney: Harcourt Brace Jovanovich Group.

Griffin, P. and Forwood, A. (1991), Adult Literacy and Numeracy Competency Scales. An International Literacy Year Project. Melbourne: Phillip Institute of Technology.

Hanna, G.S. (1985), Review of Degrees of Reading Power. *In ninth mental measurements yearbook*. Lincoln, NE: Buros Mental Measurement Institute. pp. 444-447.

Hopkins (1984), When is a Word not a Word? Riverina College of Advanced Education. Mimeo.

Hoge, R.D. (1983), *Psychometric properties of teacher-judgment measures of pupil aptitudes, classroom behaviors, and achievement levels.* Journal of Special Education. No. 17. pp. 401-429.

Hoge, R.D. and Cudmore, L. (1986), The Use of Teacher-Judgment Measures in the Identification of Gifted Pupils. *Teaching and Teacher Education*. No. 2, pp. 181-195.

Hoge, R.D. and Coladarci, T. (1989), *Teacher Based Judgements of Academic Achievement: A Review of Literature*. Review of Educational Research. Vol. 59, No. 3, pp. 297-313.

Koslin, B.L., Zeno, S., Koslin, S., Wainer, H., Ivens, S.H. (1987), The DRP: An Effectiveness Measure in Reading. New York: College Entrance Examination Board.

McNair, K. (1978-1979), Capturing Inflight Decisions: Thoughts While Teaching. *Educational Research Quarterly*. No. 3, pp. 26-42.

Mossensen, L., Hall, P. and Masters G. (1987), Tests of Reading Comprehension (TORCH). Melbourne: Australian Council for Educational Research.

National Assessment of Educational Progress (1984). The Reading Report Card. Princeton, NJ: Educational Testing Service.

Rowe, K. J. (1989), The Importance of Reading at Home. Working paper 4/89. School Programs Priorities Standing Committee. Melbourne, Vic: State Board of Education.

Stoker, H. (1988), Degrees of Reading Power. In D.J. Keyser and R.E. Sweetland (Eds), Test Critiques. Kansas City, MO: Test Corporation of America. Vol. 6.

Victoria (1990). Literacy Profiles: Teachers' Handbook. Melbourne. Ministry of Education.

Wiggins, G.P. (1989), A True Test: Toward More Authentic and Equitable Assessment. *Phi Delta Kappan*. Vol. 70, No. 9, pp. 703-713.

When is a Word Not a Word.

H. R. Hopkins. RCAE

That's an interesting question, isn't it ?
Maybe the answer lies in an understanding of what the word "word" means. But that's a pretty interesting idea too.
Y'k... k...., ..hat I mean? If ... can attach a meaning to
.. speech item ... can start playing games with it — define it
... it like that, just because it's a fun thing to do and with
change its m..n.ng by .s.ng .t w.th . d.ff.r.nt s.t .f
.th.r sp..ch .t.ms. By m.n.p.l.t.ng th. c.nt.xt w. c.n
c.ntr.l th. m..n.ng. P.rh.ps th.ts h.w w. h.v. d.v.l.p.d
control over speech and perh... ch.....n c.n un...-
st..d th. pr....ple if we g.v. th.m m.ny op...t.n
it..s to do j... th.t.

Wh.. th.. are l...n..g to b...me r..d..s, p.rh...
they need to be allowed to play with the graphic patterns
representing the spoken forms — — l.ng..g. wh.ch th..
al...dy use pr.f.c..nt...

. Maybe they sh___ l___ to r____ with a p____ of
sc.ss.rs in .n. h..d so th.. can c.t up gr.ph.c
p.t..ns and re-.ss mbl. th.m in as m___y
m_____ngf.l s.qu.nc.s as th.y c.n im.g.ne.
Th.n p..h.ps th.y c..ld d.sc.v.r ____ ans__r
to ____ f..st qu_____n j.st as y___ h.v.
d.n. by n____.

WRITING BAND B

In a nutshell ...

These students are learning about handwriting. They know what letters and words are and talk about these terms when they are writing.

CONTEXTS FOR OBSERVING BEHAVIOURS

◇ Observation of students during **WRITING SESSIONS** allows the teacher to continue to monitor increasing control and preference of hand and writing implement. Developing understandings about the conventions of print will also be apparent, and for each student, these should be noted, with dates, for records.

◇ The **WRITING CONFERENCE** is an opportunity for students to talk with teachers about their writing. This may take the form of 'reading' what has been written, or discussing the writing and/or drawing. Such times are rich sources of information and reveal much about students' understandings of the conventions and vocabulary of print and about the meaning they are making in their writing.

◇ From **SAMPLES** in the **WRITING FOLIO**, teachers will be able to identify each student's development in the use of writing conventions and the range of ideas and vocabulary in use. The **WRITING FOLIO** itself may be used for students to develop ideas for writing, as well as a list of the language conventions they are employing. This enables the teacher to locate quickly students who need assistance with ideas for writing, and also to celebrate achievements with their students.

◇ **SHARED READING** sessions are times when students are able to share their developing knowledge of the vocabulary of print, and to question those elements which are engaging their attention. Observant teachers will increase their knowledge about students during shared reading.

BEST COPY AVAILABLE

SCHOOL _____ CLASS _____ YEAR 19__

NAME _____ TERM | CONTEXT OF OBSERVATION | COMMENTS

WRITING BAND A ☐

What the writer does. ☐
Uses writing implement to make marks on paper.
Explains the meaning of marks (a word, sentence, writing, letter).
Copies "words" from signs in immediate environment.
"Reads", understands and explains own "writing".

What the writing shows ☐
Understanding of the difference between picture and print.
Use of some recognisable symbols in writing.

Use of Writing ☐
Comments on signs and other symbols in immediate environment.
Uses a mixture of drawings and "writing" to convey and support an idea.

WRITING BAND B ☐

What the writer does ☐
Reproduces words from signs and other sources in immediate environment.
Holds pencil/pen using satisfactory grip.
Uses preferred hand consistently for writing.

What the writing shows ☐
Use of vocabulary of print—letters, words, question marks.
Use of letters of the alphabet and other conventional symbols.
Use of letters in groups to form words.
Placing spaces between groups of "letters".
Knowledge that writing moves from left to right in lines from the top to the bottom of the page.

Use of Writing ☐
Writes own name.

WRITING BAND C ☐

What the writer does ☐
Commences writing without assistance.
Has a personalised handwriting style that meets most handwriting needs.
Checks written work by reading it aloud.
Sounds out words as an aid to spelling.

What the writing shows ☐
Legible writing with recognisable words.
Words put together in a sentence format.
Words written in a logical order to make a sentence that can be read.
Upper and lower case letters used conventionally.
Written sentences that can be understood by an adult.

Use of Writing ☐
Sentences convey message on one topic.
Uses "I" in writing.
Writes about feelings, judgement or direct experience.
Creates characters from experience and immediate environment.

OTHER ASSESSMENTS